Applications of the stochastic frontier approach in Energy Economics

Massimo Filippini[1] • Luis Orea[2*]

[1]*CER-ETH and CEPE, ETH Zurich and IdEP, University of Lugano, Switzerland*
[2]*Oviedo Efficiency Group, Department of Economics, University of Oviedo, Spain*

Abstract

In this paper we discuss some of the issues that energy researchers and regulators could address in two different areas of the application of stochastic frontier analysis (SFA) in Energy Economics: i) the estimation of productive or cost efficiency of electricity and gas distribution networks, and ii) the estimation of energy demand frontier models. Our review examines common problems that often appear using standard frontier models, such as the selection of input (cost) and output variables, the integration of quality incentives in cost benchmarking, the impact of distributed energy resources, or the effect of both unobserved heterogeneity and observed (e.g. environmental, socioeconomic, etc.) variables on firms' cost or energy consumption. We also point out that the SFA can be used not only to measure the level of efficiency in the use of energy, but also "rebound effects" associated with improvements in energy efficiency.

Keywords: stochastic frontier analysis, electricity and gas distribution networks, energy demand frontier models, rebound effects

JEL Classification Codes: C51, D12, D24, Q40

1. Introduction

In this paper we briefly discuss some of the issues that most researchers could address when using parametric (and often non-parametric) frontier models in Energy Economics. Due to space constraints, our review is restricted to problems that often appear using standard frontier models in two areas of application of stochastic frontier analysis (SFA): the estimation of firms' efficiency in the electricity and gas distribution sectors, and the estimation of energy demand frontier models. In recent decades the electricity and the gas distribution sectors have

---

experienced regulatory reforms with the goal of improving efficiency through the use of incentive regulation schemes such as Price-cap and Revenue-cap regulation. Therefore, the estimation of the level of productive efficiency in the distribution of electricity or gas using SFA has become relatively popular.

The second area of application, relatively new, is the estimation of the level of efficiency in the use of energy at the aggregate level. In order to reduce the environmental problems created by fossil-fuel-based energy systems, governments introduce policy instruments to promote energy efficiency improvements. These are usually examined using energy intensity ratios, such as the ratio of energy consumption to GDP. However, as discussed by Filippini and Hunt (2011), energy intensity is not a precise indicator proxy for energy efficiency because this measure is not related to the modern theory of production. The application of SFA to the analysis of energy demand not only provides policy makers with more reliable indicators of the level of efficiency in the use of energy in an economy, but also a new empirical strategy to measure the so-called "rebound effects" associated with improvements in energy efficiency.

2. Use of SFA in benchmarking regulation

In this section we briefly discuss common problems in the estimation of productive or cost efficiency in electricity distribution networks. Generally, a total cost frontier function can be specified as follows:

$$TC = F(Q, P, X, \beta) \cdot e^{v-u} \tag{1}$$

where *TC* is the total cost for each company (i.e. TOTEX); *Q* is a vector of variables representing the outputs; *P* is a vector of factor prices; *X* is a vector of output characteristics and environmental variables, including area size, quality of the service or network length, and *β* are parameters to be estimated. The total cost frontier *F(Q,P,X,β)* reflects the minimum cost for each company to produce any level of output. Unexplained variations of total cost in equation (1) are split in two parts: while *v* is the conventional noise term, *u* is a one-sided error term that is interpreted as an indicator of cost inefficiency. The most important issues to consider for the estimation of model (1) are: a) the choice of the dependent and independent variables; b) the choice of the econometric approach to be used. [1]

In electricity distribution, the variables considered as costs, inputs and outputs tend to vary significantly from one study to another (see Haney and Pollitt, 2009). For instance, some regulators only consider OPEX, leaving CAPEX outside the efficiency assessment. OPEX-oriented benchmarking should control for CAPEX if both costs are not used in fixed proportions. However, as CAPEX is likely to be an endogenous variable, benchmarking all the firms' costs (i.e. TOTEX) is often a better option.[2]

The number of customers and amount of energy distributed are the most commonly used outputs in electricity distribution. Both outputs are often highly correlated as the number of customers is the main driver of change in energy delivered (the other is per capita demand). The use of energy delivered as the unique output can be justified if per capita demand does not vary across firms or both drivers have similar effects on distribution costs. Moreover, the

---

[1] For a detailed discussion on model specification and econometric approaches used in the specification of cost functions for the distribution of energy see Farsi and Filippini (2009).

[2] It should be recalled here that some studies employing an input distance function or a DEA approach approximate the inputs using OPEX or CAPEX. This approach is only correct if input prices (in real terms) do not vary significantly across utilities. Of course, a similar bias would appear if a cost function is wrongly estimated without price variables.

increasing presence of distributed energy resources (hereafter DER) in electricity distribution may render these variables questionable as DER may simultaneously reduce the amount of energy distributed and cause additional costs. Therefore, a reduction in the outputs would paradoxically require an increase in the inputs. Indeed, Agrell and Bogetoft (2007), who include DER-related variables in their frontier models, use peak load at different voltage levels as outputs instead of the energy distributed.[3]

In designing regulatory mechanisms that consider the quality dimension of the service, regulators are faced with the task of estimating firms' marginal cost of quality improvements (MCoQ hereafter). Only Jamasb et al. (2012) and Coelli et al. (2013) have estimated MCoQ using respectively cost and distance functions. The first problem that should be addressed is that quality of service is far from being exogenously determined, as bad weather conditions tend to increase costs but also lead to lower quality services. Another key problem is the feedback between service quality and cost. Positive MCoQ are expected because quality of service is positively correlated with operating and investment practices to prevent power interruptions. However, as Jamasb et al. (2012) pointed out, the corrective costs associated with the replacement of faulty equipment might bias the estimated MCoQ downward. They also noticed that preventative and corrective costs maintain a different relationship with quality of service over time, and hence MCoQ can be measured from a cost derivative with respect to expected values of quality of service.

An additional issue in benchmarking regulation has to do with the effect of environmental factors on network costs (see Yu et al., 2009; Growitsch et al., 2012). Controlling for environmental effects is challenging as the number of environmental variables should be kept down in order to avoid dimensionality problems. In this sense, the use of popular statistical variable reduction techniques may be useful. Orea et al. (2012) advocate using supervised dimension reduction methods (e.g. sliced inverse regression, SIR) that have been extensively applied in other scientific fields in recent years.

From the econometric point of view, the estimation of electricity cost functions is mostly carried out using panel data. Several panel data estimators have been proposed and used for this purpose. Of particular interest are the clustering and non-clustering methods recently developed to deal with unobserved heterogeneity across firms.[4] The non-clustering approach includes the panel data frontier models introduced by Greene (2005), where unobserved heterogeneity is captured through a set of firm-specific intercepts (more details can be found in Section 3). The clustering approach can be viewed as a two-stage procedure where the original sample is first split into a number of mutually exclusive groups (classes), and then separate efficiency analyses are carried out for each class/sub-sample. Thus, the clustering approach allows the estimation of different technological characteristics for firms belonging to different groups. In this sense, Agrell et al. (2013) advocate using the so-called latent class model to split the sample as this approach is specifically designed to cluster firms by searching for differences in production or cost parameters. Moreover, the simulation analyses carried out by Llorca et al. (2014) show that the latent class approach outperforms other somewhat more arbitrarily and less robust procedures of splitting a sample of observations, such as cluster analysis.[5]

---

[3] Cossent (2013) also points out that it is unclear how to compute the number of consumers in a situation where DER can be connected under net-metering schemes.

[4] Some studies argue that discrepancies in efficiency estimates between different approaches can be due to the presence of unobserved heterogeneity. See Farsi and Filippini (2004) and Farsi et al. (2006) for a discussion of this issue.

[5] Dai and Kuosmanen (2013) pointed out that this approach is also becoming quite popular because it applies well to the conventional DEA as well as to the StoNED method. See, for instance, Po et al. (2009).

3. Use of SFA in energy demand models

In this second part of the paper, we summarize a new approach proposed by Filippini and Hunt (2011) on the empirical measurement of the "underlying energy efficiency". This approach, based on the stochastic frontier analysis, treats energy as a production factor used in combination with other inputs to produce energy services.[6] The analysis of the level of "underlying energy efficiency" is generally based on the econometric estimation of an aggregate energy demand frontier function that can be viewed as the reduced-form model of an underlying structural model based on a household utility optimizing problem.[7] This energy demand frontier function can be estimated for the whole economy or just for a sector, such as the residential sector. Generally, an aggregate energy demand frontier function is specified as follows:

$$E = f(Y, P, X, t, \beta) \cdot e^{v-u} \tag{2}$$

where E is the aggregate energy consumption; Y is income; P is the real energy price; β are parameters to be estimated; X is a vector of variables influencing demand for energy, including climate, characteristics of the household, structure of the economy, etc.; and t is a time trend capturing both 'exogenous' technical progress and other common factors.[8] Unexplained variations of energy demand are again captured by two random terms, v and u. Whereas the noise term v in (2) is assumed to be normally distributed, the non-negative random term u is generally assumed to follow a half-normal distribution. The energy demand frontier $f(Y, P, X, t, \beta)$ reflects the minimum amount of energy necessary to produce any given level of energy service (Filippini and Hunt 2011, 2013). Therefore, the positive deviations from the energy demand frontier function captured by u can be interpreted as inefficiency in the use of input energy.

Filippini and Hunt (2011, 2012) proposed to use panel data and a relatively simple log-log functional form. Regardless of the specification of the energy demand frontier, equation (2) implicitly assumes that responses to changes in both income and energy prices do not depend on energy efficiency. This implies that the energy demand function is separable. If this separability condition is not satisfied but still we impose it, we could bias the estimated elasticities and energy efficiency scores. In an energy demand setting, this seems to be a quite restrictive assumption because more energy-inefficient consumers should have more elastic energy demands (as they are more penalized when the price of energy increases).[9]

The estimation of the stochastic energy demand frontier function (2) can be performed using cross-sectional data or panel data. The use of panel data stochastic frontier models is important because of the presence of unobserved heterogeneity. The estimation of (2) can be therefore carried out using a number of different SFA model specifications such as the pooled model, the random effects model (REM), the true fixed effects model (TFEM), and the true random effects model (TREM).[10] The TREM and the TFEM are non-clustering methods to control for unobserved variables and extend the original pooled model to panel data by adding an individual effect. Therefore, these models consider two stochastic terms for unobserved heterogeneity, one for time-variant unobserved factors (captured by the noise term) and one

---

[6] Some studies have been published on the measurement of the energy efficiency using a Data Envelopment Approach (DEA). See, for example, Hu and Wang (2006) and Wei et al. (2009).

[7] See, for instance, Flaig (1990) and Alberini and Filippini (2011).

[8] As suggested in Filippini and Hunt (2012), time dummies can also be used to capture a common Underling Energy Demand Trend (UEDT).

[9] Consumers' stickiness or bounded rationality might yield however less elastic demands.

[10] For a general presentation of these models, see Greene (2008) and Farsi and Filippini (2009).

for the unit-specific time-invariant socioeconomic and environmental characteristics (captured by the individual effects).[11]

Orea et al. (2014) bring attention to the fact that the above energy demand frontier approach is closely connected to the measurement of the so-called "rebound effect" associated with improvements in energy efficiency.[12] The rebound effect refers to the reduction in the unit cost of energy services (e.g. heating, lighting and refrigeration) when the energy system becomes more efficient. This price reduction can generate an increase in the consumption of the energy service that might offset the expected savings in energy consumption provided by an energy efficiency improvement.

In this sense, these authors point out that as equation (2) includes an energy inefficiency term, the elasticity of demand for energy with respect to changes in this term (i.e. $\varepsilon_E = -\partial \ln E / \partial u$) provides a direct measure of the rebound effect. The actual saving in energy consumption will only be equal to the predicted saving from engineering calculations when $\varepsilon_E = -1$, and hence there is no rebound effect.[13] The rebound effect would be positive if actual savings in energy consumption are less than expected, i.e. $-1 < \varepsilon_E$. However, since ε_E in (2) is equal to -1, the traditional SFA energy demand frontier model implicitly imposes a zero rebound effect, which contradicts most of the available empirical evidence.

Orea et al. (2014) relax the zero-rebound effect assumption and suggest estimating the following energy demand frontier model:

$$\ln E = \ln f(Y, P, X, t, \beta) + v + g[ES(z, \gamma)] \cdot u \qquad (3)$$

where *g*[.] is simply the rebound effect, and *γ* are new parameters to be estimated. Several interesting remarks could be made regarding the measurement of rebound effects with a SFA model. First, the rebound effect function *g* depends on the demand of energy services *ES*, which like the energy inefficiency level is unobserved. This latent variable is approximated by *z,* a set of energy services determinants (income and price variables). Second, the above model can be viewed as a traditional heteroscedastic SFA model in production economics that satisfies the traditional scaling property.[14] However, while the (scaling) function *g* is interpreted in the latter literature as a portion of total estimated inefficiency, Orea et al. (2014) treat the (rebound effect) function *g* as part of the deterministic frontier. Third, equation (3) precludes the existence of negative rebound effects or 'backfire' outcomes as *g≥0* should be imposed in order to distinguish inefficiency from noise. Finally, the traditional exponential modelling of *g* might yield 'over conservation' rebounds effects for some observations (see Saunders, 2008). One strategy to deal with this somewhat counterintuitive outcome is using an index-type function that lies between zero and one (see, for instance, the scaling function proposed by Kumbhakar, 1990).

---

[11] As discussed by Farsi et al. (2005), the REM and the TREM can suffer from the "unobserved variables bias". In order to solve the unobserved heterogeneity bias in the estimation of stochastic frontier models, these authors proposed to use the Mundlak (1978) adjustment. This adjustment is based on an auxiliary equation with the group-means of the explanatory variables.

[12] For a detailed summary of this literature, see Sorrell and Dimitropoulos (2008).

[13] Here we use the definition provided by Saunders (2000). Following this author, the rebound effect can be measured as R=1+ε_E.

[14] See Álvarez *et al.* (2006) for a review of the models that incorporate this property in the literature on frontier production functions.

4. Further issues and challenges for future research

In the preceding sections we have discussed a selection of empirical issues that often appear when standard SFA models are used in both benchmarking regulation and energy demand settings. Obviously, there are other interesting challenges facing applied researchers in benchmarking regulation such as the estimation of semi-parametric stochastic frontier models, the estimation of transient and persistent efficiency, and the estimation of dynamic frontier models. Further research is also desirable to find potential feedbacks between frontier models and engineering-oriented models, such as reference network models and norm models. Regarding the estimation of energy consumption, we could also consider issues such as measurement errors in prices and the identification of asymmetric price effects, the control for unobserved heterogeneity using panel data estimators when (most of the) variables change slowly over time and the estimation of frontier models with endogenous variables. Finally, additional research is also needed in order to find flexible (but tractable) rebound-effect functions that are able to depict the full range of rebound values (including backfire outcomes).

References

Agrell, P. and Bogetoft, P. (2007) *Development of benchmarking models for German electricity and gas distribution*, Project GERNER-Final report, SUMICSID, 2007-01-01.

Agrell, P.J., Farsi, M., Filippini, M. and Koller, M. (2013) *Unobserved heterogeneous effects in the cost efficiency analysis of electricity distribution systems*, CER-ETH Economics working papers series 13/171, CER-ETH - Center of Economic Research (CER-ETH) at ETH Zurich.

Alberini, A. and Filippini, M. (2011) Response of residential electricity demand to price: the effect of measurement error, *Energy Economics*, 33(5), 889–895.

Álvarez, A., Amsler, C., Orea, L., and Schmidt, P. (2006) Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics, *Journal of Productivity Analysis*, 25, 201-212.

Coelli, T., Gautier, A., Perelman, S. and Saplacan-Pop, R. (2013) Estimating the cost of improving quality in electricity distribution: a parametric distance function approach, *Energy Policy*, 53, 287–329.

Cossent, R. (2013) *Economic regulation of distribution system operators and its adaptation to the penetration of distributed energy resources and smart grid technologies*, PhD project, Escuela Técnica Superior de Ingeniería (ICAI), mimeo.

Dai, X. and Kuosmanen, T. (2013) Best-practice benchmarking using clustering methods: application to energy regulation, *Omega*, 42, 179–218.

Farsi, M. and Filippini, M. (2004) Regulation and measuring cost efficiency with panel data models application to electricity distribution utilities, *Review of Industrial Organization,* 25(1), 1-19.

Farsi, M. and Filippini, M. (2009) *Efficiency measurement in the electricity and gas distribution sectors*, in Evans, J. and Hunt, L.C. (eds.): International Handbook on the Economics of Energy, Edward Elgar, Cheltenham, 598-623.

Farsi, M., Filippini, M. and Greene, W. (2005) Efficiency measurement in network industries: Application to the Swiss railway companies, *Journal of Regulatory Economics*, 28, 69-90.

Farsi, M., M. Filippini and W. H. Greene (2006) Application of panel data models in benchmarking analysis of the electricity distribution sector, *Annals of Public and Cooperative Economics*, 77(3), 271-290.

Filippini, M. and Hunt, L. (2011) Energy demand and energy efficiency in the OECD countries: A stochastic demand frontier approach, *Energy Journal*, 32(2), 59-80.

Filippini, M. and Hunt, L. (2012) US residential energy demand and energy efficiency: A stochastic demand frontier approach, *Energy Economics*, 34, 1484-1491.

Filippini, M. and Hunt, L. (2013) *Underlying energy efficiency' in the US*, CER-ETH Economics Working Paper Series 13/181.

Flaig, G. (1990) Household production and the short-and long-run demand for electricity, *Energy Economics*, 12(2), 116–121.

Greene, W.H. (2005) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model, *Journal of Econometrics*, 126, 269-303.

Greene, W.H. (2008) *The econometric approach to efficiency analysis*, in H.O. Fried, C.A.K. Lovell and S.S. Schmidt (eds.): The measurement of productive efficiency and productivity growth, Oxford University Press, 92–250.

Growitsch, C., Jamasb, T. and Wetzel, H. (2012) Efficiency effects of observed and unobserved heterogeneity: Evidence from Norwegian electricity distribution networks, *Energy Economics*, 34, 542–548.

Haney, A.B. and Pollitt, M. (2009) Efficiency analysis of energy networks: An international survey of regulators, *Energy Policy*, 37(12), 5814–5830.

Hu, J. and Wang, S. (2006) Total-factor energy efficiency of regions in China, *Energy Policy*, 34, 3206-3217.

Jamasb, T., Orea, L. and Pollitt, M. (2012) Estimating marginal cost of quality improvements: The case of the UK electricity distribution companies, *Energy Economics*, 34, 1498–1506.

Kumbhakar, S. (1990) Production frontiers, panel data, and time-varying technical efficiency, *Journal of Econometrics*, 46, 201–211.

Llorca, M., Orea, L. and Pollitt, M. (2014) Using the latent class approach to cluster firms in benchmarking: an application to the US electricity transmission industry, *Operations Research Perspectives*, 1(1), 6-17.

Mundlak, Y. (1978) On the pooling of time series and cross section data, *Econometrica*, 64, 69-85.

Orea, L., Growitsch, C. and Jamasb, T. (2012) *Using supervised environmental composites in production and efficiency analyses: An application to Norwegian electricity networks*, EWI Working Paper 12/18, Institute of Energy Economics, University of Cologne.

Orea, L., Llorca, M. and Filippini, M. (2014) *Measuring energy efficiency and rebound effects using a stochastic demand frontier approach: the US residential demand for energy*, Efficiency Series Paper 01/2014, Oviedo Efficiency Group, Department of Economics, University of Oviedo.

Po, R., Guh, Y. and Yang, M. (2009) A new clustering approach using data envelopment analysis, *European Journal of Operational Research*, 199, 276-284.

Saunders, H. (2000) A view from the macro side: Rebound, backfire and Khazzoom–Brookes, *Energy Policy*, 28 (6–7), 439–449.

Saunders, H. (2008) Fuel conserving (and using) production functions, *Energy Economics*, 30(5), 2184-2235.

Sorrell, S. and Dimitriopolous, J. (2008) The rebound effect: microeconomic definitions, limitations and extensions, *Ecological Economics*, 65(3), 636-649.

Wei, C., Ni, J. and Shen, M. (2009) Empirical analysis of provincial energy efficiency in China, *China and World Economy*, 17(5), 88-103.

Yu, W., Jamasb, T. and Pollitt, M. (2009) Does weather explain cost and quality performance? An analysis of UK electricity distribution companies, *Energy Policy*, 37(11), 4177-4188.